UCRL-ID-153499

DNA Compaction by Yeast Mitochondrial Protein ABF2p

R.W. Friddle, J.E. Klare, A. Noy, M. Corzett, R. Balhorn, R.J. Baskin, S.S. Martin, E.P. Baldwin

May 9, 2003

U.S. Department of Energy

Lawrence
Livermore
National
Laboratory

DISCLAIMER

This work was performed under the auspices of the U. S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.

DNA Compaction by Yeast Mitochondrial Protein ABF2p

Raymond W. Friddle, Jennifer E. Klare, and Aleksandr Noy*

Biosecurity and Nanoscience Laboratory,

Chemistry and Materials Science Directorate,

Lawrence Livermore National Laboratory,

7000 East Ave, Livemore, CA 94550

Michelle Corzett and Rodney Balhorn

Biology and Biotechnology Program,

Lawrence Livermore National Laboratory,

7000 East Ave, Livermore, CA 94550

Ronald J. Baskin, Shelley S. Martin, and Enoch P. Baldwin

Department of Molecular and Cellular Biology,

University of California at Davis, Davis, CA 95616

Abstract

We used high resolution Atomic Force Microscopy (AFM) to image compaction of linear and circular DNA by the yeast mitochondrial protein ABF2p , which plays a major role in maintaining mitochondrial DNA. AFM images show that protein binding induces drastic bends in the DNA backbone for both linear and circular DNA. At high concentration of ABF2p DNA collapses into a tight globular structure. We quantified the compaction of linear DNA by measuring the end-to-end distance of the DNA molecule at increasing concentrations of ABF2p. We also derived a polymer statistical mechanics model that gives quantitative description of compaction observed in our experiments. This model shows that a number of sharp bends in the DNA backbone is often sufficient to cause DNA compaction. Comparison of our model with the experimental data showed excellent quantitative correlation and allowed us to determine binding characteristics for ABF2. Our studies indicate that ABF2 compacts DNA through a novel mechanism that involves bending of DNA backbone. We discuss the implications of such a mechanism for mitochondrial DNA maintenance.

*Correspondence should be addressed to A.N.: noy1@llnl.gov

INTRODUCTION

Mitochondria participate in many critical processes in the cell lifecycle. Aside from its primary role in ATP production, mitochondria also act as signaling centers through regulation of calcium, iron and metabolite levels in the cytosol. These organelles are also responsible for the main switch controlling apoptosis. Such critical responsibilities place stringent requirements on the integrity of the mitochondrial DNA (mtDNA). A variety of processes threatens mtDNA. Respiratory chain of mitochondrial metabolism produces large levels of oxygen radicals which can attack mtDNA. Oxidative damage to mtDNA often leads to several clinical disorders including Parkinson's, Hutchinson's, and Huntington's diseases [1]. Ironically, the very job that is required to keep the cell alive also yields dangerous byproducts. In order to operate under these harsh conditions mitochondria need to package mtDNA in a way that protects it from damage, while not impairing the normal functions of mtDNA such as replication and transcription.

Mammals [2, 3] and the budding yeast *S. cerevisiae* [4–6] package mtDNA in compact globular structures similar to a bacterial nucleoid. These *mt-nucleoid* structures are distinctly different from the packaging of DNA into chromatin in the cell nucleus. Researchers have firmly established the mechanism of histone proteins action in packaging of nuclear DNA. However, very little data exists on the identity or function of the proteins that facilitate the formation of the mt-nucleoid.

Diffley and Stillman found that a particular 20 kDa protein was present in relatively high abundance among the various polypeptides isolated from mt-nucleoids. This protein, ABF2p (ARS binding factor 2) displays interesting DNA binding characteristics: it binds non-specifically to general regions of DNA, but exhibits phased binding to replicating sequences such as ARS1 [7]. ABF2p also induces negative supercoiling in DNA in the presence of topoisomerase. While ABF2p is not required for mtDNA replication,changes in ABF2p levels alter mtDNA copy number [8], and null ABF2p mutants lose their wild type (ρ^+) mtDNA [9]. Data also indicate that levels of ABF2p di-

rectly influence the number of recombination intermediates *in vivo* [10]. This information, coupled with the high abundance of ABF2p , lead researchers to suggest that ABF2p is a primary mt-DNA packaging protein. ABF2p is closely related to the HMG family; its sequence contains two HMG boxes linked by six amino acids [9]. HMG proteins are, among other activities, involved in the structural organization of packaged DNA in higher ordered structures such as chromatin. However no known DNA packaging mechanism uses HMG proteins as the *fundamental* packaging unit. Therefore, it is likely that the complete DNA packaging mechanism employed by mitochondria is different from other known DNA packaging processes. The establishment of such mechanism should provide valuable information about the role of ABF2p in the overall mt-DNA maintenance process.

Recent developments in molecular-scale imaging powered a number of unprecedented advances in biophysics. For example, high resolution atomic force microscopy (AFM) can now visualize single biological molecules in native environments in real space [11]. Researchers have used AFM to image protein binding to DNA [**?**], virus particle surfaces, cell surfaces and determine the strength of protein-ligand interactions and elasticity of DNA molecule [**?** **?**]. AFM also excels in visualizing the conformation of linear polymers such as DNA [12–14]. Early on, Balhorn and colleagues used AFM to investigate the mechanism of the DNA compaction by protamine proteins from sperm [15]. Others have used AFM to study the physical properties of DNA condensation by various ionic species which may in fact have biological significance in terms of the mechanism used by viruses to package DNA [16–18]. Recent refinements in AFM imaging technology, such as new imaging modes and new sharper AFM probes [**?** **?**], pushed the technique's limits even further.

In this paper we investigated binding of ABF2p to DNA using high resolution atomic force microscopy. We found that when ABF2p bound to DNA it induced pronounced structural distortions in DNA conformation. When we increased the protein coverage we observed a striking collapse of the DNA molecule into a dense globular complex. Our observations lead us to suggest that ABF2p uses a novel mechanism to compact DNA simply by introducing

a number of sharp bends into the DNA backbone. This mechanism is distinctly different from all previously known mechanisms of DNA compaction. We also present a simple mathematical model that describes the compaction process quantitatively and we compare this model with the experimental observations. Finally, we used our model to derive ABF2p binding parameters from the AFM data.

MATERIALS AND METHODS

ABF2p Isolation The pMALc2X vector (New England Biolabs) containing residues 21-177 of the ABF2 gene was inserted between the initiator Met and the six His residues of the pET28b(+) expression vector (Novagen). The pET28b-ABF2 vector was introduced into *E. coli* BL21(DE3) for expression. Purification of the His_6 tagged protein was carried out initially with 500 mM NaCl lysis buffer, then dialyzed to low salt buffer prior to ion exchange chromatography. Concentrated protein samples at 20-70 mg/ml (20 mM Na-HEPES pH 7.5, 1 mM Na-EDTA, 4nM DTT, 0.1 w/v Na-azide), were then diluted to 4 mg/ml in Tris buffer (10 mM Tris-Cl, pH 7.8, 4 mM DTT, 1 mM Na-EDTA). ABF2p binding activity to either pLitmus-38(+) or linear lambda phage DNA (New England Biolabs) was assayed by gel-retardation.

pBR322 DNA Relaxed circular pBR322 (TopoGen) was used in AFM images of circular DNA. Linear DNA was prepared by digest of supercoiled pBR322 at the BamH1 site.

AFM Imaging and Analysis Substrates were prepared for imaging by first applying 3 μl of 0.1% poly-L-lysine (PL) to freshly cleaved mica. After 1 minute the mica was rinsed with copious amounts of water and dried with filtered N_2 stream for 1 minute. DNA/ABF2p complexes were prepared in 100 mM $NaHCO_3$ buffer. DNA concentration was held constant at 1 μg/ml and ABF2p concentrations were varied. After we mixed the protein and DNA solutions we allowed them to equilibrate for 5 minutes; then 1 μl of sample solution was applied to the PL coated mica substrate and allowed to adsorb for 1 - 2 minutes at room temperature. The substrate was then rinsed with

water and finally dried with filtered N_2 stream. All images were acquired in air using a Multimode Nanoscope IIIa Atomic Force Microscope (Digital Instruments). Scans were taken in tapping mode using etched silicon FESP tips (NanoWorld).

Measurements of DNA end-to-end distances and protein-induced bend angles were done using image analysis tools built into the Nanoscope IIIa software. We assume the DNA are pulled to the PL-coated mica surface with little reorganization of the DNA strands in the plane parallel to the surface. Therefore we take the images to represent the 2-dimensional projection of the molecule's native 3-dimensional configuration in solution. The end-to-end distances of the DNA therefore represent a projection of the true distance and to compare them with our 3-dimensional model we must multiply our average measured distances $\langle R \rangle_m$ by a factor of $\pi/2$ (see Appendix for details).

We fir our corrected average measured end-to-end distances to our model using the Levenberg-Marquardt algorithm to carry out a non-linear χ^2-minimization in IGOR Pro 4.0 (Wavemetrics).

Circular Dichroism Linear DNA (pBR322) in 100 mM $NaHCO_3$ at $\sim$50 μg/ml were titrated with aliquots of concentrated ABF2p in a 1 cm path length quartz cuvette. CD spectra were taken using a J-715 spectropolarimeter (Jasco) operated at room temperature. We first subtracted the baseline from the raw CD spectra and then smoothed them using a 2^{nd} order, 11 point Savitzky-Golay algorithm in IGOR Pro 4.0 (Wavemetrics). For each protein concentration $[\text{ABF2p}]_i$, the protein CD at 275 nm was subtracted from the corresponding DNA/protein CD.

AFM IMAGING OF ABF2P-DNA

To investigate protein binding to DNA we imaged individual pBR322 plasmid DNA molecules after we incubated them with ABF2p . After incubating the protein and DNA in solution, we immobilized DNA-protein complexes on atomically-flat mica surfaces pre-treated with poly-L-lysin. We kept the DNA concentration in solution low enough to obtain single isolated molecules on the

surface. AFM images of the DNA molecules before exposure to protein show smooth contours of the backbone, typically devoid of sharp bends or other distinguishing features (Figure 1A). In contrast, when we exposed DNA to the ABF2p we observed multiple sharp bends in the DNA backbone. Images at low concentration of the ABF2p at most instances showed a protein molecule bound at the site of the bend (at least to within the limits of AFM resolution). Therefore we can conclude that binding of the protein results in the formation of these bends. This behavior is not entirely surprising, since other HMG family proteins often induce structural distortions in the DNA [19–22]. As we increased the concentration of the protein, the number of bends in the DNA backbone also increased. At higher concentrations of ABF2p we clearly started to observe the overall compaction of the molecule(Figure 1C). Finally, at very high concentration of the protein DNA collapsed into globular nucleoid-like structures (Figure 1D,E).

Diffley and Stillman reported that ABF2 induced negative supercoiling in DNA [9]. It is unclear, however, if supercoiling plays a significant role in the compaction mechanism. To test this possibility we imaged ABF2p complexes with relaxed circular DNA. We used relaxed circular pBR322 plasmid for these experiments, which allowed us to eliminate any possibility of sequence dependence influencing our experiments. Remarkably, AFM images of ABF2p complexed with circular DNA show almost identical behavior to what we observed for linear DNA. Protein binding induced large bends in the DNA backbone and the number of bends progressively increased with the increase in the protein concentration. At high ABF2p concentration DNA again collapsed into globular structures. This experiment strongly indicates that supercoiling does not play a significant role in the mechanism of the ABF2p actions. Rather, we believe that negative supercoiling observed by Diffley and Stillman can simply be a byproduct of the structural distortions that ABF2p introduces into the backbone of circular DNA.

We can also use AFM images of the partially-compacted DNA at low concentrations of ABF2p to estimate the angle of the DNA backbone bend induced by the protein. We have measured the bending angle for 43 different DNA

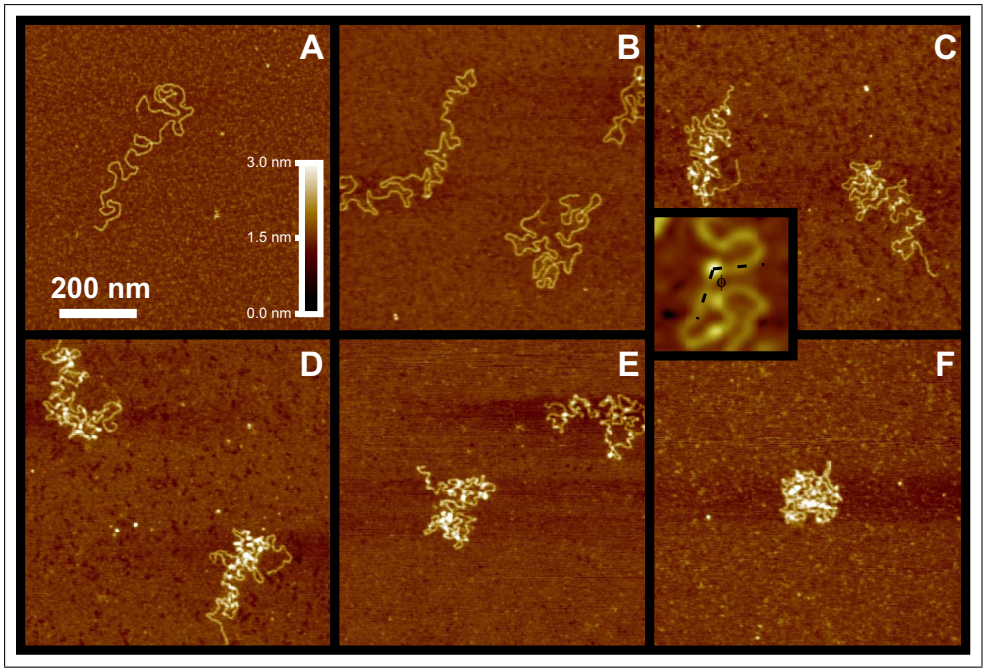

FIG. 1: AFM images of linearized pBR322 DNA molecules after exposure to increasing concentrations of ABF2p .**A**: No ABF2p ; **B** 1.5 μg/mL; **C** 3.5 μg/mL; **D** 7 μg/mL;**E** 15 μg/mL; **D** 25 μg/mL; Inset: Close-up image of a bend in the DNA backbone induced by the bound protein (bright dot).

molecules which had isolated protein units bound to DNA (Figure 1, Inset). Histogram of the bend angles shows a peak at 102° (Figure 3). This value is clearly different from the mica lattice angle of 120°; therefore, we are confident that mica does not cause this structural distortion. The bends caused by the ABF2p occur over a very short distance, which also distinguishes them from the random bends that sometimes appear when DNA adsorbs to the surface. Therefore our images indicate that ABF2p bends DNA by about 78° (if we adopt the angle measuring convention common to structural studies). Our measured value of the bend angle compare favorably with the literature values for the bend angles that other proteins from the HMG family induce in DNA. For example, SRY bends DNA by about 70° [23] and SOX bends DNA by about 83° [24]. However, a direct comparison between the literature values and our measured bend angle value is difficult since all the reported structures contain only one HMG box, whereas ABF2p contains two HMG boxes. At

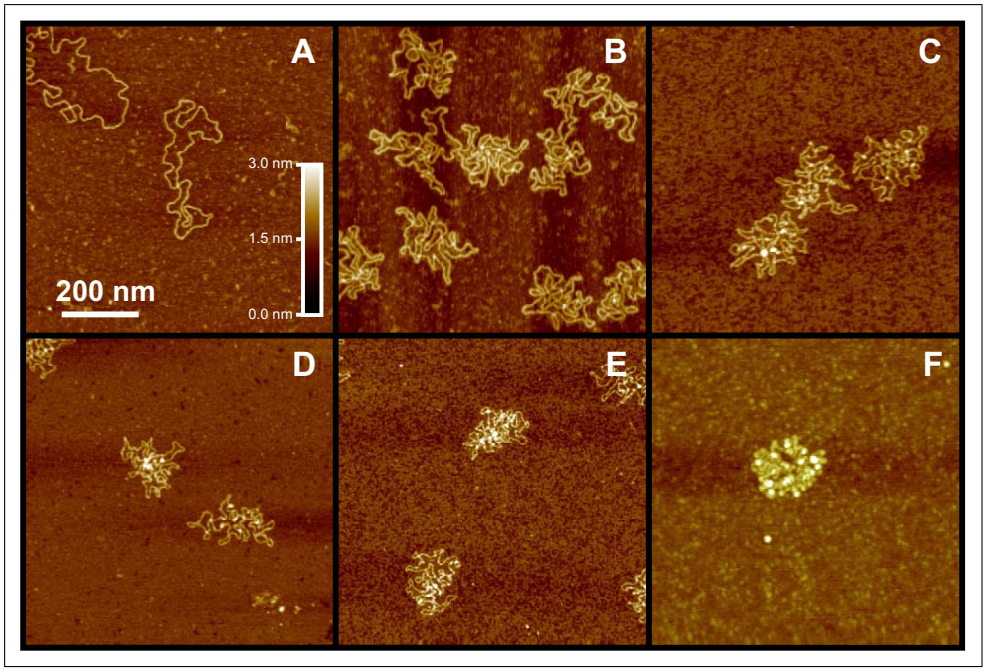

FIG. 2: AFM images of relaxed circular pBR322 DNA molecules after exposure to increasing concentrations of ABF2p .**A**: No ABF2p ; **B** 1.5 μg/mL; **C** 3.5 μg/mL; **D** 7 μg/mL; **E** 15 μg/mL; **D** 25 μg/mL.

present, resolution constraints do not allow us to determine whether ABF2p binds to the DNA as a monomer or as a dimer, or whether the bend that we observe consists of two closely spaced sub-bends. Further investigations utilizing higher resolution probes, such as single-wall carbon nanotube AFM tips [25, 26] might provide more information about the binding mode for ABF2p .

Our AFM images do not show any further type of aggregation or distortion besides the bend formation up until very high concentrations of protein, when it becomes difficult to discern individual protein molecules in the tightly compacted globular structure. We also have not found any clear evidence of cooperative binding. Mostly, the protein-DNA complexes showed random distribution of bends throughout the length of DNA, whereas cooperative binding would have forced a clustering of the bends in certain regions of DNA. Our images suggest that the major effect of the ABF2p binding to the DNA is the formation of sharp bends in the DNA backbone. Qualitatively, such bends would reduce the intrinsic stiffness of DNA and lead to the overall reduction of

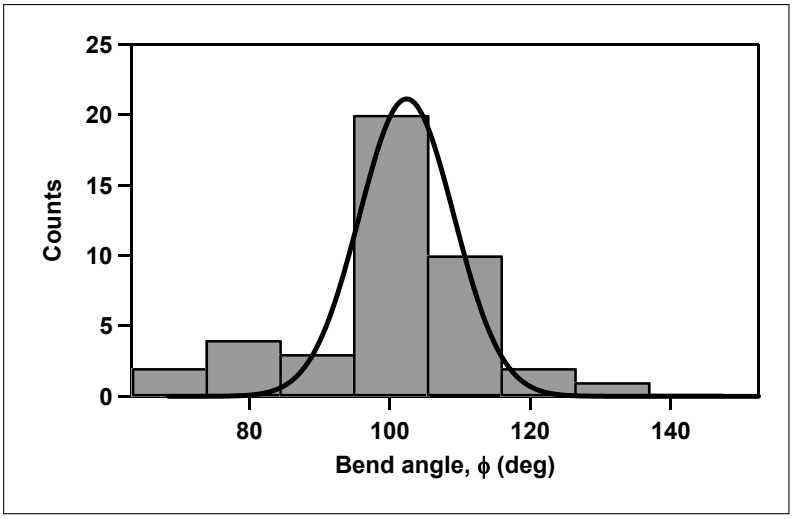

FIG. 3: Histogram of the bend angles measured from 42 images of individual DNA molecules exposed to ABF2p

the DNA size in solution. However, it is unclear whether bending of the DNA backbone alone is sufficient to induce compaction. To answer this question quantitatively we need to consider how such bending changes the dynamics of a stiff polymer chain in solution. To address this question we developed a model that accounts for small, localized distortions in the DNA backbone while maintaining the nature of a DNA molecule elsewhere along its contour.

DNA-protein complex conformation: *bent*-Worm-Like Chain Model

We base our model on Kratky and Porod's Worm-Like Chain (WLC) model [27] which describes the statistical behavior of a random flight polymer chain that also has an intrinsic stiffness associated with it. We modify this model to include a number P of fixed bends of angle ϕ. For simplicity we assume that these bends are uniformly distributed along the DNA helix. We then derive the expression for the mean-squared end-to-end distance for such a polymer chain. We include a detailed mathematical derivation of this *bent*-Worm-Like Chain ($bWLC$) model in the Appendix, and here we just state the main premises and results of this model.

We start with a DNA chain of contour length L. The intrinsic stiffness of

the DNA is governed by its persistence length A, which incorporates effects of temperature, charge, screening, and solute-solvent interactions. We assume that the bends that we introduce into the DNA do not affect its structure in the regions between the bends. In addition, we allow the dihedral angles to rotate freely in our model. We obtain the following exact analytical solution for the mean-squared end-to-end distance for an ensemble of linear polymers that incorporate P equidistant bends of angle ϕ:

$$\begin{aligned}
\left\langle R^2 \right\rangle_{bWLC} &= \left[\frac{1+P-(P+2)\Gamma+\Gamma^{P+2}}{(1-\Gamma)^2}\right]\left[\frac{2AL}{P+1}-2A^2\left(1-e^{-\frac{L}{A(P+1)}}\right)\right] \\
&- \left[\frac{P\Gamma-(P+1)\Gamma+\Gamma^{2(P+1)}}{(1-\Gamma)^2}\right]\left[\frac{2AL}{P+1}+2A^2\left(1-e^{\frac{L}{A(P+1)}}\right)\right] \quad (1) \\
\Gamma &= -\cos\phi\; e^{-\frac{L}{A(P+1)}}
\end{aligned}$$

This result is consistent with the original Kratky-Porod model since it recovers the original WLC expression when no bends are present ($P=0$). The most interesting aspect of this expression is that it predicts a significant drop in the average end-to-end distance as the number of bends increases (Figure). This effect is most noticeable when the bend angle approaches 90° and it diminishes as the bend angles approach 180°. When the bend angle becomes equal to 180°, the bends vanish and we again recover the original WLC model behavior. Significantly, our model shows that the introduction of bends into the DNA backbone can alone cause significant compaction of the DNA. Therefore, it is consistent with the mechanism of ABF2p action that we propose.

As compaction proceeds the frequency of DNA-DNA interactions will increase. Therefore contributions to the global DNA structure from the steric repulsion between DNA segments will increase with added protein. We expect that in the last stages of compaction when the segments are very close together the contribution from these excluded volume effects should become rather significant. Since we are representing DNA as a phantom chain we do not take into account any of the steric repulsion effects. To make our model

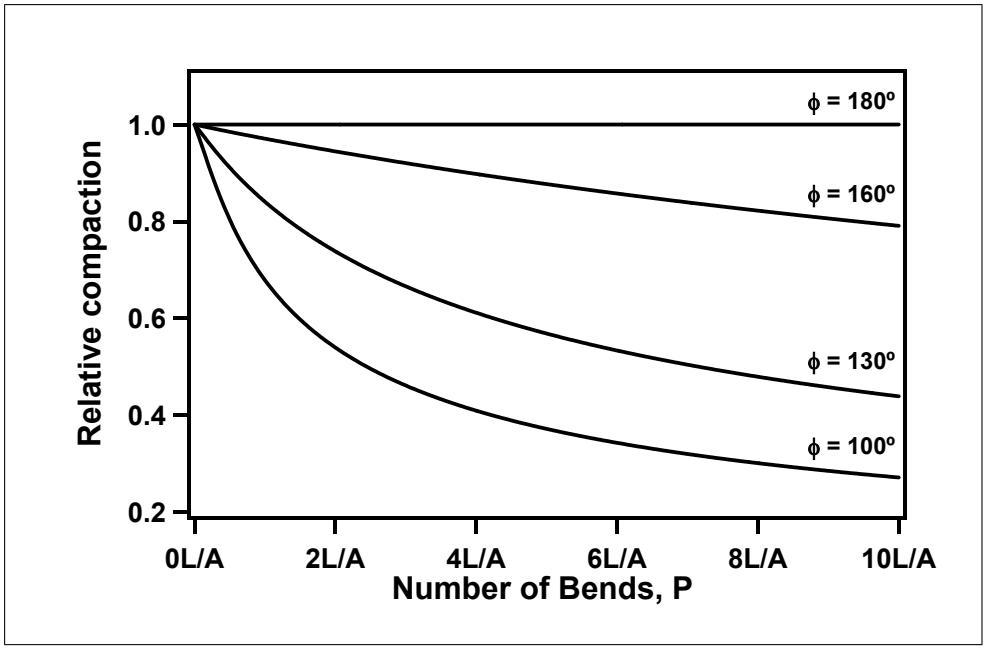

FIG. 4: b-WLC model predictions for DNA compaction as a function of the number of bends. We define compaction as decrease in he average end-to-end distance $\sqrt{< R^2 >}$ relative to the average end-to-end distance for the free DNA.

more realistic we need to introduce a correction that would account for these excluded volume effects. Treatment of DNA segments interacting with one another can become quite complex if we attempt to account explicitly for charge repulsion, Leonard-Jones potential, salt counter-ion shielding, etc. Moreover, the interaction becomes virtually impossible to calculate if we try to include bound proteins in the overall potential. However it has been shown that such detailed descriptions of the pair potential is unnecessary given that the short range potential is conserved [28, 29] Therefore, it is reasonable to neglect long range potentials and simply approximate the interaction between segments as a short range repulsive force [30]. Several authors have applied perturbation theory to this problem and obtain the following first order correction [30–35]:

$$\begin{aligned} \left\langle R^2 \right\rangle &\cong \left\langle R^2 \right\rangle_o \left(1 + \frac{4}{3} z + \cdots\right) \\ z &= \left(\frac{3}{2\pi \left\langle R^2 \right\rangle_o}\right)^{3/2} \beta n^2 \end{aligned} \tag{2}$$

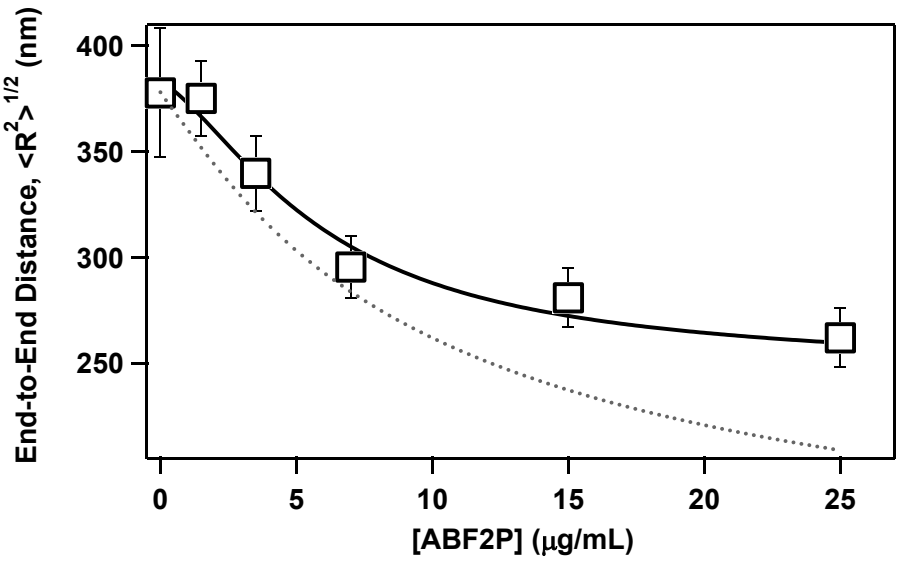

FIG. 5: Plot of average end-to-end distance for linear pBR322 DNA molecules and function of average ABF2p concentration. Lines represent the b-WLC model fits to the data without (dashed line) and with (solid line) excluded volume correction.

where $\langle R^2 \rangle_o$ is the mean-squared end-to-end distance in the unperturbed state, n is the number of Kuhn statistical segments, and β is the binary cluster integral for a pair of segments. The parameter β represents the effective volume excluded to one segment by another. Odijk and other estimated the excluded volume between two cylindrical Kuhn segments of length b and diameter D as $\pi b^2 D/2$ [36, 37]. As the protein binds to the double helix it increases the effective diameter of the cylinder from the 2 nm value characteristic of the free DNA. TO calculate the excluded volume correction in this case we assumed that we can represent the areas of DNA covered with protein as helices of larger diameter and that the contribution of these larger helices is proportional to the amount of the bound protein (see Appendix for details).

We tested our model by comparing average end-to-end distances obtained in the experiment with the model predictions. AFM images of individual compacted DNA molecules give us the ability to measure the end-to-end distance directly for a large number of molecules and collect necessary statistics. Figure 5 shows that the measured end-to-end distance sharply decreases with the

increase in the ABF2p concentration, as expected from images on Figure 1. We can fit our data to the b-WLC model if we assume that (a) the number of bends,P, is proportional to the amount of bound protein, and (b) ABF2p binding can be described with a simple binding constant K_d and a Hill constant q.

$$P = P_{max} \frac{(C/K_d)^q}{1 + (C/K_d)^q} \tag{3}$$

where P_{max} is the maximum number of bends that a protein can create in a given DNA length. We used the 102° value for the bend angle that we measured in our experiments and assumed that the maximum number of bends in the DNA is determined by the pBR322 length and ABF2p footprint. If we use the literature value of 30 base pairs as ABF2p footprint [9], then the maximum number of bends (i.e. binding sites) that the protein can introduce into pBR322 DNA (which contains 4,361 base pairs) is 145. When we fit our experimental data to our model using these values we obtained an excellent fit to the experimental data (Figure 5, solid line). For comparison, the b-WLC model without the excluded volume correction provided a much less satisfactory fit (Figure 5, dotted line), especially at high protein concentration where steric repulsion makes the greatest contribution. We obtained the best fit to the experimental data for the protein binding constant $K_D = 1.38\,\mu$M and Hill coefficient of $n = 1.2$. The model fit also yields the value of the persistence length for free DNA of 48 nm, which matches the values found by experiments using the standard WLC model [38, 39]. These comparisons indicate that our model captures the essential physics of the compaction process induced by ABF2p .

Finally, to compare the values of protein dissociation K_D predicted by our model with the experimental K_D value we used CD spectroscopy to study binding of ABF2p to DNA. The technique is based on that used by McAfee et. al. in the study of equilibrium binding of Sac7d and Sso7d to DNA [40, 41]. It has been shown that DNA CD in the region 250-310 nm is sensitive to conformational changes in the double helix [42, 43]. Conveniently, this band generally has the weakest CD for protein spectra. Therefore, this

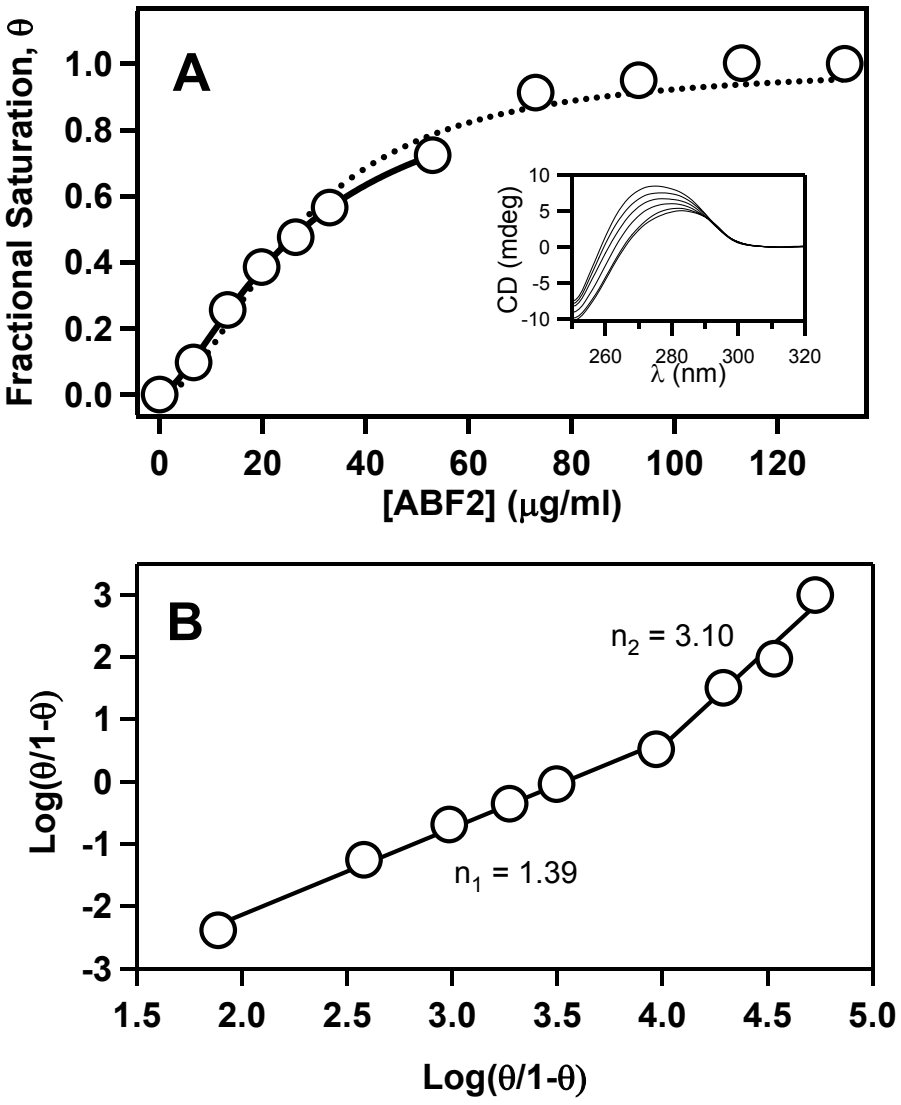

FIG. 6: **A.** Plot of fraction of bound linearized pBR322 DNA molecules as a function of ABF2p concentration as determined from the CD spectra. Dashed line represents the fit to the whole data set to a cooperative binding model. Solid line represents to only the initial region of the Hill plot. Inset shows the CD spectra. **B** Hill plot of the data in part A.

region represents a good choice for monitoring binding of a protein that distorts DNA, such as ABF2p. CD spectra obtained at increasing concentrations of ABF2p clearly showed an increase in DNA distortion that saturates at protein concentrations above $80\,\mu$g/mL (Figure 6A). The same data plotted

in Hill coordinates (Figure 6B) indicate that ABF2p binding is weakly cooperative with the Hill constant of 1.4. The Hill plot shows that at ABF2p concentrations above 60 μg/mL the cooperativity parameter sharply rises to 3.1; yet, we need to interpret these data with caution. As some literature examples show, rather than indicate the switch in protein binding mechanism this change might indicate a change in the structure of the protein-DNA complex that would have a stronger effect on the CD properties [?]. Significantly, the CD data for the protein binding in the initial regime (Figure 6A, solid line) yield the binding constant $K_D = 1.4\,\mu$M, which is almost identical to the value of $K_D = 1.38\,\mu$M that we obtained from our AFM data and the b-WLC model. The values of the Hill coefficient that we obtained from the AFM images (n=1.2) and the CD data (n=1.4) also show very good agreement. This comparison validates the b-WLC model and also confirms our hypothesis that ABF2p compacts DNA simply by placing sharp bends along the DNA double helix.

This mechanism can lead only to vaguely defined globular conformations of the DNA-protein complexes, which is quite similar to the appearance of the mt-nucleoid. It is also important to note that ABF2p action is strikingly different from other known compaction mechanisms, which typically involve much more ordered superstructures such as chromatin in the case of histone-induced packaging in the nucleus, or DNA toroids formes by protamines in sperm cells.

Further studies would be necessary to uncover the role of such a packaging arrangement in the regulation of DNA in mitochondria. However, we can speculate that such loose packing could simplify access of various regulatory proteins to DNA. Such an arrangement could allow the mitochondrion to avoid the need for a sophisticated DNA handling apparatus similar to the multi-stage machinery present in the eucaryotic cell nuclei.

CONCLUSIONS

We have used high resolution atomic force microscopy to observe binding of mitochondrial protein ABF2p to linear and circular DNA plasmids. Surprisingly, our images showed that protein biding induces sharp bends of about 102° in the DNA backbone. AFM images also clearly showed that ABF2p binding results in the compaction of DNA molecules for both linear and circular DNA. Moreover, at high protein concentration DNA molecules collapsed into compact globular structures reminiscent of a mitochondrial nucleoid. To analyze this compaction process we have developed a statistical model that describes the DNA-protein complex as a stiff polymer chain that has sharp bends incorporated throughout its length. Using this model, we have shown that incorporation of bends into the DNA backbone is alone sufficient to cause DNA compaction. End-to-end distances predicted by the model showed excellent agreement with the end-to-end distances that we measured from the AFM images. Moreover, binding parameters that we obtained from our model showed excellent agreement with the results of bulk studies. Significantly, the ABF2p compaction mechanism that we established appears to be distinctly different from common DNA packaging proteins.

Our findings have important implications for several areas. First, we showed that high-resolution AFM imaging can provide quantitative information regarding protein-DNA interactions. Single molecule imaging not only can provide information on the geometrical conformation of the protein-DNA complexes, but also can determine thermodynamic binding parameters. We believe that AFM will be an important part of the modern biophysics toolkit for studies of protein-DNA interactions. Second, we believe that our results will be important for establishing the mechanisms of mitochondrial DNA maintenance and regulation. The apparent loose packing of DNA by the ABF2p should provide important clues for the structure of the mitochondrial nucleoid and for possible access pathways for regulatory proteins. Further AFM studies using other proteins should provide a wealth of information about maintenance and regulation of mitochondrial DNA.

Acknowledgements

RWF is supported by the LLNL SEGRF fellowship. JEK acknowledges support from AWU and LLNL. This work was performed under the auspices of U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under contract #W 7405-Eng-48.

---

[1] D. Voet, J. G. Voet, C. W. Pratt, *Fundamentals of biochemistry*, (Wiley, New York, 1999).

[2] G. C. Van Tuyle, M. L. McPherson, *J Biol Chem* **254**, 6044 (1979).

[3] M. Satoh, T. Kuroiwa, *Experimental Cell Research* **196**, 137 (1991).

[4] U. Wintersberger, M. Binder, P. Fischer, *Mol Gen Genet* **142**, 13 (1975).

[5] I. Miyakawa, H. Aoi, N. Sando, T. Kuroiwa, *J Cell Sci* **66**, 21 (1984).

[6] I. Miyakawa, N. Sando, S. Kawano, S. Nakamura, T. Kuroiwa, *J Cell Sci* **88 (Pt 4)**, 431 (1987).

[7] J. F. Diffley, B. Stillman, *Proc Natl Acad Sci U S A* **85**, 2120 (1988).

[8] O. Zelenaya-Troitskaya, S. M. Newman, K. Okamoto, P. S. Perlman, R. A. Butow, *Genetics* **148**, 1763 (1998).

[9] J. F. Diffley, B. Stillman, *Proc Natl Acad Sci U S A* **88**, 7864 (1991).

[10] D. M. MacAlpine, P. S. Perlman, R. A. Butow, *Proc Natl Acad Sci U S A* **95**, 6739 (1998).

[11] H. G. Hansma, K. J. Kim, D. E. Laney, R. A. Garcia, M. Argaman, M. J. Allen, S. M. Parsons, *J Struct Biol* **119**, 99 (1997).

[12] Y. L. Lyubchenko, A. A. Gall, L. S. Shlyakhtenko, *Methods Mol Biol* **148**, 569 (2001).

[13] Y. L. Lyubchenko, L. S. Shlyakhtenko, M. Binus, C. Gaillard, F. Strauss, *Nucleic Acids Res* **30**, 4902 (2002).

[14] S. Tiner, W. J., V. N. Potaman, R. R. Sinden, Y. L. Lyubchenko, *J Mol Biol* **314**, 353 (2001).

[15] M. J. Allen, E. M. Bradbury, R. Balhorn, *Nucleic Acids Res* **25**, 2221 (1997).

[16] J. H. Hoh, Y. Fang, *Molecular Biology of the Cell* **10**, 179a (1999). Meeting Abstract.

[17] Y. Fang, J. H. Hoh, *FEBS Letters* **459**, 173 (1999).

[18] Y. Fang, J. H. Hoh, *Journal of the American Chemical Society* **120**, 8903 (1998).

[19] D. N. Jones, M. A. Searles, G. L. Shaw, M. E. Churchill, S. S. Ner, J. Keeler, A. A. Travers, D. Neuhaus, *Structure* **2**, 609 (1994).

[20] J. E. Masse, B. Wong, Y. M. Yen, F. H. Allain, R. C. Johnson, J. Feigon, *J Mol Biol* **323**, 263 (2002).

[21] M. H. Werner, M. E. Bianchi, A. M. Gronenborn, G. M. Clore, *Biochemistry* **34**, 11998 (1995).

[22] J. J. Love, X. Li, D. A. Case, K. Giese, R. Grosschedl, P. E. Wright, *Nature* **376**, 791 (1995).

[23] Y. Tang, L. Nilsson, *Proteins-Structure Function and Genetics* **31**, 417 (1998).

[24] P. Scaffidi, M. Bianchi, *J. Bio. Chem.* **276**, 47296 (2001).

[25] G. R. Schnitzler, C. L. Cheung, J. H. Hafner, A. J. Saurin, R. E. Kingston, C. M. Lieber, *Mol Cell Biol* **21**, 8504 (2001).

[26] C. L. Cheung, J. H. Hafner, C. M. Lieber, *Proc Natl Acad Sci U S A* **97**, 3809 (2000).

[27] O. Kratky, G. Porod, *Rec. Trav. Chim. Pays-Bas.* **68**, 1106 (1949).

[28] B. H. Zimm, *J. Chem. Phys.* **14**.

[29] W. H. Stockmayer, *Macromol. Chem.* **35**, 54 (1960).

[30] H. Yamakawa, *Modern theory of polymer solutions*, (Harper Row, New York,, 1971).

[31] T. B. Grimley, *J. Chem. Phys.* **21**, 185 (1953).

[32] T. B. Grimley, *J. Chem. Phys.* **22**, 1134 (1954).

[33] T. B. Grimley, *Trans. Faraday Soc.* **55**, 687 (1959).

[34] B. H. Zimm, W. H. Stockmayer, M. Fixman, *J. Chem. Phys.* **21**, 1716 (1953).

[35] N. Saito, *J. Phys. Soc. Japan* **9**, 780 (1954).

[36] T. Odijk.

[37] L. Onsager, *Ann. N. Y. Acad. Sci.* **51**, 627 (1949).

[38] C. Bustamante, J. F. Marko, E. D. Siggia, S. Smith, *Science* **265**, 1599 (1994).

[39] C. G. Baumann, S. B. Smith, V. A. Bloomfield, C. Bustamante, *Proc Natl Acad Sci U S A* **94**, 6185 (1997).

[40] J. G. McAfee, S. P. Edmondson, I. Zegar, J. W. Shriver, *Biochemistry* **35**, 4034 (1996).

[41] S. P. Edmondson, J. W. Shriver, *Methods Enzymol* **334**, 129 (2001).

[42] W. A. Baase, J. Johnson, W. C., *Nucleic Acids Res* **6**, 797 (1979).

[43] B. B. Johnson, K. S. Dahl, J. Tinoco, I., V. I. Ivanov, V. B. Zhurkin, *Biochemistry* **20**, 73 (1981).

[44] P. J. Flory, *Statistical mechanics of chain molecules*, (Interscience Publishers, New York,, 1969). also reprinted by Hanser publishers, 1989.

APPENDIX

$\langle \mathbf{R}^2 \rangle_{bWLC}$ Worm-Like Chain with Bends

Consider a piece of a freely rotating chain (FRC) with some arbitrary number of segments p (for a detailed explanation of the FRC see P.J. Flory [44] or Yamakawa [30]). We join several of these pieces together such that the joint between pieces makes an angle ϕ. All other joint angles within the pieces are equal to θ. Rotation about each bond is unrestricted. Therefore $\theta_i = \theta$ for $i = 1, 2, \ldots, p-1$ and $\theta_p = \phi$ for each piece in the chain. The mean squared end-to-end distance of the chain is given by,

$$\begin{aligned} \langle \mathbf{R}^2 \rangle &= \sum_{i=1}^{n} \sum_{j=1}^{n} \langle \mathbf{r}_i \cdot \mathbf{r}_j \rangle \\ &= \sum_{i=1}^{n} \langle \mathbf{r}_i^2 \rangle + 2 \sum_{i=1}^{n-1} \sum_{j>i}^{n} \langle \mathbf{r}_i \cdot \mathbf{r}_j \rangle \end{aligned} \tag{4}$$

The first sum in equation (4) is simply the sum over a^2. The second sum is a little more involved. The factor of 2 ensures that we need only take the product of segment $i = 1$ with $j = 2, \ldots, n$ and segment $i = 2$ with $j = 3, \ldots, n$ and so on up to $i = n - 1$ with $j = n$. Consider the case of $i = 1$, we will look at one piece from $j = 2, \ldots, p$.

$$\sum_{j=2}^{p} \langle \mathbf{r}_1 \cdot \mathbf{r}_j \rangle = \sum_{j=2}^{p} \alpha^{j-1}$$

Summing over the next piece should include the fixed angle ϕ in place of a θ. Let $\beta = -\cos\phi$, then for $j = p+1, \ldots, 2p$ we have,

$$\sum_{j=p+2}^{2p} \frac{\beta}{\alpha} \alpha^{j-1} = \alpha^p \frac{\beta}{\alpha} \sum_{j=2}^{p} \alpha^{j-1}$$

Likewise, the sum over the third piece of the chain must include two factors of β and remove two factors of α,

$$\sum_{j=2p+2}^{3p} \frac{\beta^2}{\alpha^2} \alpha^{j-1} = \alpha^{2p} \left(\frac{\beta}{\alpha} \right)^2 \sum_{j=2}^{p} \alpha^{j-1}$$

We can see the $(m+1)^{th}$ piece is given by,

$$\alpha^{mp}\left(\frac{\beta}{\alpha}\right)^m \sum_{j=2}^{p} \alpha^{j-1}$$

So for a chain of n segments divided into $\frac{n}{p}$ pieces, there will be $\frac{n}{p}-1$ bends. The entire sum of projections of $j > i$ for $i = 1$ is therefore given by,

$$\sum_{j=2}^{n} \langle \mathbf{r}_1 \cdot \mathbf{r}_j \rangle = a^2 \sum_{m=0}^{\frac{n}{p}-1} \alpha^{mp}\left(\frac{\beta}{\alpha}\right)^m \sum_{j=2}^{p} \alpha^{j-1} \tag{5}$$

Notice in equation (5) that the starting index in the sum over j is $j = 2$. Therefore we lose the projection of segment $\mathbf{r}_{p+1}$ on to $\mathbf{r}_1$ for each peice in the chain (ie. the first segment of each peice is not included in the sum over j). Likewise, when $i = 2$ we lose the first and second segments of each piece. These lost projections will be accounted for in other sums later. Let (5) inserted into the second term of (4) be called $\mathbb{A}$, not to be confused with the persistence length A. So to multiply every segment in the first piece with the rest of the $\frac{n}{p} - 1$ pieces (neglecting lost terms) we have,

$$\mathbb{A}_0 = \sum_{m=0}^{\frac{n}{p}-1} \alpha^{mp}\left(\frac{\beta}{\alpha}\right)^m \sum_{i=1}^{p-1} \sum_{j>i}^{p} \alpha^{j-1}$$

The second piece is identical to the first except we now need only sum over $\frac{n}{p} - 2$ pieces,

$$\mathbb{A}_1 = \sum_{m=0}^{\frac{n}{p}-2} \alpha^{mp}\left(\frac{\beta}{\alpha}\right)^m \sum_{i=1}^{p-1} \sum_{j>i}^{p} \alpha^{j-1}$$

and so on up until we reach the last piece in which it simply sums over itself and we only need $m = 0$. Therefore we write the entire sum as,

$$\mathbb{A} = \sum_{s=0}^{\frac{n}{p}-1} \mathbb{A}_s = \sum_{s=0}^{\frac{n}{p}-1} \sum_{m=0}^{\frac{n}{p}-1-s} \alpha^{mp}\left(\frac{\beta}{\alpha}\right)^m \sum_{i=1}^{p-1} \sum_{j>i}^{p} \alpha^{j-i} \tag{6}$$

Equation (6) is just a combination of geometric series that can be easily solved. The result is given by,

$$\mathbb{A} = \left[\frac{\frac{n}{p} - \left(\frac{n}{p}+1\right)Q + Q^{\frac{n}{p}+1}}{(1-Q)^2}\right]\left[\frac{p\alpha}{1-\alpha} - \frac{\alpha - \alpha^{n+1}}{(1-\alpha)^2}\right] \tag{7}$$

where $Q = \alpha^{p-1}\beta$. Now we will account for the lost terms in the sum (6). As mentioned before, the construction of $\mathbb{A}$ periodically skips over segments as it sums over i. By inspection of (6) we can see the following:

for i=	we miss j=
1	mp+1
2	mp+1, mp+2
⋮	⋮
p-1	mp+1, mp+2, …, mp+p-1

where $m = 1, 2, \ldots, \frac{n}{p} - 1$. We therefore construct the following sum which contains the above terms,

$$\mathbb{B} = \sum_{s=0}^{\frac{n}{p}-2}\sum_{m=1}^{\frac{n}{p}-1-s} \alpha^{mp}\left(\frac{\beta}{\alpha}\right)^m \sum_{i=1}^{p-1}\sum_{j=1}^{i}\alpha^{j-i} \tag{8}$$

The solution can again be simplified into geometric series which result in,

$$\mathbb{B} = \left[\frac{(\frac{n}{p}-1)Q - \frac{n}{p}Q^2 + Q^{2\frac{n}{p}}}{(1-Q)^2}\right]\left[-\frac{\alpha^2}{(1-\alpha)^2}(1-\alpha^{-p}) - p\frac{\alpha}{1-\alpha}\right] \tag{9}$$

A third and final term is needed to make up for another missed product in the sum (6). This is to account for the dot product of the p^{th} segment of each piece with all higher number segments. This sum is simply given by,

$$\mathbb{C} = \sum_{s=0}^{\frac{n}{p}-2}\sum_{m=0}^{\frac{n}{p}-2-s} \alpha^{mp}\frac{\beta^{m+1}}{\alpha^m}\sum_{j=1}^{p}\alpha^{j-1} \tag{10}$$

Which can be solved similarly as $\mathbb{A}$ and $\mathbb{B}$. The result is,

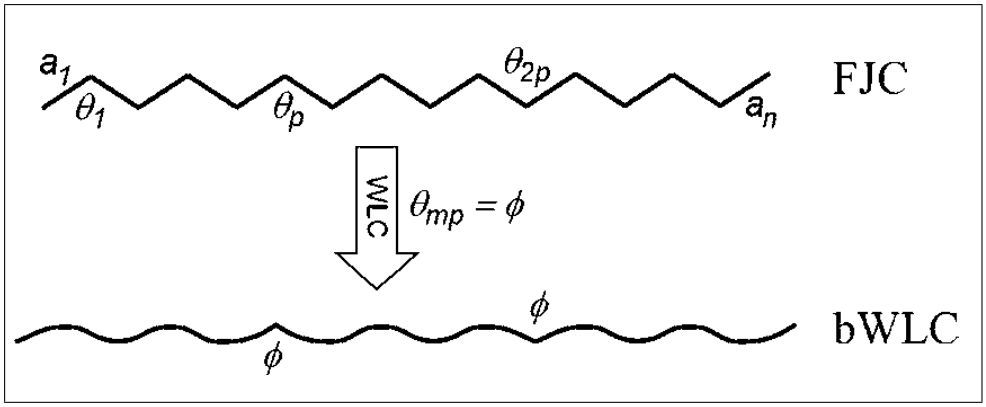

FIG. 7: Schematic representation of the relationship between the b-WLC model and the freely-rotating chain (FRC) model. We start with a FRC model that contains abnormal bend angles at regular intervals. The final b-WLC model retains the discontinuous bends of angle ϕ, yet the polymer chain is smooth everywhere else.

$$\mathbb{C} = \left[\frac{Q^{\frac{n}{p}}\beta - Q\beta + (\frac{n}{p} - 1)\beta}{(1-Q)^2}\right]\left[\frac{1-\alpha^p}{1-\alpha}\right] \tag{11}$$

Now the terms $\mathbb{A}$, $\mathbb{B}$, and $\mathbb{C}$ comprise the double sum in (4). So inserting these terms we have,

$$\langle \mathbf{R}^2 \rangle = na^2 + 2a^2\big[\mathbb{A} + \mathbb{B} + \mathbb{C}\big] \tag{12}$$

At this point, equation (12) describes a freely rotating chain with *abnormal* bends at an angle ϕ that are positioned equidistant from one another along the chain. We call these abnormal because they are different from the normal bends θ found throughout in the regular FRC. The interval distance between abnormal bends is inversely proportional to the number of these bends since the overall contour length of the chain L must remain constant. (Note: Yamakawa [30] gives a general formulation for a FRC with periodic structure. Although we did not use this in our derviation of (12) it presumably could be used to arrive at the same result.)

Finally, we take the limit to the WLC by the well known procedure employed by Kratky and Porod [27, 44] to generate a semi-stiff continuous contour chain (Figure 7). The following constraints are held constant in the limit,

$$\lim_{a\to 0,\ \alpha\to 1} \frac{a}{1-\alpha} \equiv A, \qquad (13)$$
$$\lim_{a\to 0,\ n\to\infty} na \equiv L$$

where A is the persistence length derived from the normal FRC. For the present model we add another constraint for the length of each piece in the chain. This is,

$$\lim_{a\to 0,\ p\to\infty} pa \equiv T$$

here T is the length between adjacent bends. Making the estimation that as $\alpha \to 1$ then $-\ln\alpha \approx 1-\alpha$ and solving (13) for α we have,

$$\alpha = e^{-a/A} \qquad (14)$$

Now, using the above constraints we can simplify each term in $\langle \mathbf{R}^2 \rangle$ of equation (12).

$$\begin{aligned}
2a^2\mathbb{A} &= \left[\frac{\frac{L}{T} - (\frac{L}{T}+1)e^{-T/A}\beta + \left(e^{-T/A}\beta\right)^{\frac{L}{T}+1}}{\left(1-e^{-T/A}\beta\right)^2}\right]\left[2TA - 2A^2\left(1-e^{-T/A}\right)\right] \\
2a^2\mathbb{B} &= \left[\frac{(\frac{L}{T}-1)e^{-T/A}\beta - \frac{L}{T}\left(e^{-T/A}\beta\right)^2 + \left(e^{-T/A}\beta\right)^{2\frac{L}{T}}}{\left(1-e^{-T/A}\beta\right)^2}\right]\left[-2TA - 2A^2\left(1-e^{T/A}\right)\right] \\
2a^2\mathbb{C} &= 2aA\left[\frac{e^{-\frac{L}{A}}\beta^2 - e^{-T/A}\beta^2 + (\frac{L}{T}-1)\beta}{\left(1-e^{-T/A}\beta\right)^2}\right]\left(1-e^{-T/A}\right) \\
&\overset{a\to 0}{=} 0
\end{aligned}$$

The first term in equation (12) is zero since $na^2 = La = 0$. The complete function for $\langle \mathbf{R}^2 \rangle_{bWLC}$ in terms of the length between bends T is given by,

$$\begin{aligned}
\langle \mathbf{R}^2 \rangle_{bWLC} = &\left[\frac{\frac{L}{T} - (\frac{L}{T}+1)e^{-T/A}\beta + \left(e^{-T/A}\beta\right)^{\frac{L}{T}+1}}{\left(1-e^{-T/A}\beta\right)^2}\right]\left[2TA - 2A^2\left(1-e^{-T/A}\right)\right] \\
+ &\left[\frac{(\frac{L}{T}-1)e^{-T/A}\beta - \frac{L}{T}\left(e^{-T/A}\beta\right)^2 + \left(e^{-T/A}\beta\right)^{2\frac{L}{T}}}{\left(1-e^{-T/A}\beta\right)^2}\right]\left[-2TA - 2A^2\left(1-e^{T/A}\right)\right]
\end{aligned}$$

A more useful form of this function is found by noticing that the number of bends, call it P, is given by $P = \frac{L}{T} - 1$. This is the form given in equation (1) where substituting $T = \frac{L}{P+1}$ gives $\langle \mathbf{R}^2 \rangle_{bWLC}$ in terms of the number of bends.

Excluded Volume

Due to increased segment density the interaction between DNA segments is enhanced with compaction. We therefore look to relate the parameters in the first order perturbation theory of equation (2) to the number of bends. As bends are induced in the chain, some regions are perturbed more so than others. Only at high coverage of the DNA will the entire chain have uniform interaction parameters describing its state. Therefore for intermediate numbers of bends we will approximate the situation with average parameters linearly related to the number of bends. We now define an average binary cluster integral,

$$\langle\beta\rangle = \frac{P\beta_l + (P_{max} - P)\beta_s}{P_{max}} \tag{15}$$

where β_s and β_l represent the unperturbed *segment* and the superhelical *loop* cluster integrals respectively. As these superhelical loops form, the effective contour length of the chain is reduced. We therefore write the number of segments as a function of increasing bends in the following way:

$$n_o = \frac{n_l - n_s}{P_{max}} P + n_s \tag{16}$$

where n_s and n_l are the number of segments in the free chain and completely covered chain respectively.

Assuming ABF2p binds as a monomer, Diffley and Stillman show the DNA footprint to be approximately 30 bp. Therefore, assuming a 30 bp$\times$0.34 nm/bp = 10 nm arc segment we estimate the radius of one superhelical loop as $r = 10\text{nm}/(102^\circ\pi/180^\circ)$ and therefore the diameter of the superhelical Kuhn segments is given by $D_l = 2r = 11.2$ nm. AFM measurements yield an estimate to the superhelical pitch of 15 nm. Using three ABF2 molecules per loop ($3 \times 102^\circ = 306^\circ \sim 360^\circ$) and using pBR322 DNA with 145 binding sites/3 = 48 loops, we have a final effective contour length of $n_l = 48 \times 15$ nm = 725 nm when completely covered in protein. With these calculations we find the following parameters for ABF2p binding to linearized pBR322,

$$
\begin{aligned}
\beta_s &= \frac{\pi(100\text{ nm})^2 2\text{ nm}}{2} = 31.3 \times 10^3\text{ nm}^3 \\
\beta_l &= \frac{\pi(100\text{ nm})^2 11.2\text{ nm}}{2} = 175.9 \times 10^3\text{ nm}^3 \\
n_s &= 4361\text{ bp}\frac{0.34\text{ nm/bp}}{100\text{ nm}} = 14.8 \\
n_l &= 48\text{ loops}\frac{15\text{ nm/loop}}{100\text{ nm}} = 7.25
\end{aligned}
$$

Finally, the expression for the mean-squared end-to-end distance with excluded volume correction is then given by:

$$
\begin{aligned}
\langle R^2 \rangle &= \langle R^2 \rangle_{bWLC} \left(1 + \frac{4}{3}z\right) \\
z &= \left(\frac{3}{2\pi\langle R^2 \rangle_{bWLC}}\right)^{3/2} \langle\beta\rangle n_o^2
\end{aligned}
$$

2D Correction Factor

Call the angle between a vector from end-to-end and the plane of the surface θ. Since we are not interested in direction, we can ignore the azimuthal angle and simply consider possible orientations from $\theta = 0$ (parallel to surface) and $\theta = \pi/2$ (orthogonal to surface). Averaging over possible projections we estimate that, on average, the projected $\langle R \rangle_m$ and real $\langle R \rangle$ vectors differ by a factor $\pi/2$.

$$
\begin{aligned}
\langle R \rangle_m &\approx \langle R \rangle \langle \cos\theta \rangle \\
\langle \cos\theta \rangle_0^{\pi/2} &= \frac{2}{\pi} \\
\therefore \quad \langle R \rangle &\approx \frac{\pi}{2}\langle R \rangle_m
\end{aligned}
$$